DOCUMENT RESUME

ED 470 205 TM 034 345

AUTHOR	Osborne, Jason W.; Waters, Elaine
TITLE	Multiple Regression Assumptions. ERIC Digest.
INSTITUTION	ERIC Clearinghouse on Assessment and Evaluation, College Park, MD.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	EDO-TM-02-05
PUB DATE	2002-08-00
NOTE	4p.
CONTRACT	ED-99-CO-0032
AVAILABLE FROM	ERIC Clearinghouse on Assessment and Evaluation, 1129 Shriver Laboratory, University of Maryland, College Park, MD 20742. Tel: 800-464-3742 (Toll Free); Web site: http://ericae.net.
PUB TYPE	ERIC Publications (071) -- ERIC Digests in Full Text (073)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Error of Measurement; *Nonparametric Statistics; *Regression (Statistics); *Reliability
IDENTIFIERS	ERIC Digests; Homoscedasticity (Statistics); Linear Models; *Normal Distribution; Type I Errors; *Violation of Assumptions

ABSTRACT

This Digest presents a discussion of the assumptions of multiple regression that is tailored to the practicing researcher. The focus is on the assumptions of multiple regression that are not robust to violation, and that researchers can deal with if violated. Assumptions of normality, linearity, reliability of measurement, and homoscedasticity are considered. Checking these assumptions carries significant benefits for the researcher, and making sure an analysis meets the associated assumptions helps avoid Type I and II errors. Attending to such issues as attenuation due to low reliability, curvilinearity, and nonnormality often boosts effect sizes, usually a desirable outcome. There are many nonparametric statistical techniques available to researchers when the assumptions of a parametric statistical technique are not met. These are often somewhat lower in power than parametric techniques, but they provide valuable alternatives for researchers. (SLD)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 470 205

Multiple Regression Assumptions

Jason W. Osborne
North Carolina State University

Elaine Waters
University of Oklahoma

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM034345

ERIC Digest EDO-TM-02-05 August 2002

BEST COPY AVAILABLE

Multiple Regression Assumptions

Jason W. Osborne and Elaine Waters
North Carolina State University and University of Oklahoma

Most statistical tests rely upon certain assumptions about the variables used in the analysis. When these assumptions are not met, the results may not be trustworthy, resulting in a Type I or Type II error, or over- or underestimation of significance or effect size(s). As Pedhazur (1997, p. 33) notes, "Knowledge and understanding of the situations when violations of assumptions lead to serious biases, and when they are of little consequence, are essential to meaningful data analysis." However, as Osborne, Christensen, and Gunter (2001) observe, few articles report having tested assumptions of the statistical tests they rely on for drawing their conclusions. This creates a situation in which we have a rich literature in education and social science, but we are forced to call into question the validity of many of these results, conclusions, and assertions, as we have no idea whether the assumptions of the statistical tests were met. Our goal for this Digest is to present a discussion of the assumptions of multiple regression tailored toward the practicing researcher.

Several assumptions of multiple regression are "robust" to violation (e.g., normal distribution of errors), and others are fulfilled in the proper design of a study (e.g., independence of observations). Therefore, we will focus on the assumptions of multiple regression that are not robust to violation, and that researchers can deal with if violated. Specifically, we will discuss the assumptions of normality, linearity, reliability of measurement, and homoscedasticity.

Normality Assumption

Regression assumes that variables have normal distributions. Non-normally distributed variables (highly skewed or kurtotic variables, or variables with substantial outliers) can distort relationships and significance tests. There are several pieces of information that are useful to the researcher in testing this assumption: visual inspection of data plots, skew, kurtosis, and P-P plots give researchers information about normality, and Kolmogorov-Smirnov tests provide inferential statistics on normality. Outliers can be identified either through visual inspection of histograms or frequency distributions, or by converting data to z-scores.

Bivariate/multivariate data cleaning can also be important (Tabachnick & Fidell, 2001, p 139) in multiple regression. Most regression or multivariate statistics texts (e.g., Pedhazur, 1997; Tabachnick & Fidell, 2001) discuss the examination of standardized or studentized residuals, or indices of leverage. The removal of univariate and bivariate outliers can reduce the probability of Type I and Type II errors and improve accuracy of estimates.

Outlier (univariate or bivariate) removal is straightforward in most statistical software. However, it is not always desirable to remove outliers. In this case, transformations (e.g., square root, log, or inverse) can improve normality but complicate the interpretation of the results and should be used deliberately and in an informed manner. A full treatment of transformations is beyond the scope of this Digest, but is discussed in many popular statistical textbooks.

Linearity Assumption

Standard multiple regression can only accurately estimate the relationship between dependent and independent variables if the relationships are linear in nature. Because there are many instances in the social sciences in which nonlinear relationships occur (e.g., anxiety), it is essential to examine analyses for nonlinearity. If the relationship between independent variables (IV) and the dependent variable (DV) is not linear, the results of the regression analysis will underestimate the true relationship. This underestimation carries two risks: increased chance of a Type II error for that IV, and, in the case of multiple regression, an increased risk of Type I errors (overestimation) for other IVs that share variance with that IV.

Authors such as Pedhazur (1997), Cohen and Cohen (1983), and Berry and Feldman (1985) suggest three primary ways to detect nonlinearity. The first method is to use theory or previous research to inform current analyses. However, because many prior researchers have probably overlooked the possibility of nonlinear relationships, this method is not foolproof. A preferable method of detection is to examine residual plots (plots of the standardized residuals as a function of standardized predicted values, readily available in most statistical software).

The third method of detecting curvilinearity is to routinely run regression analyses that incorporate curvilinear components (squared and cubic terms; see Goldfeld and Quandt, 1976 or most regression texts for details on how to do this) or use the nonlinear regression option available in many statistical packages. It is important that the nonlinear aspects of the relationship be accounted for in order to best assess the relationship between variables.

Reliability Assumption

The nature of our educational and social science research means that many variables we are interested in are also difficult to measure, making measurement error a particular concern. In simple correlation and regression, unreliable measurement causes relationships to be *underestimated*, increasing the risk of Type II errors. In the case of multiple regression or partial correlation, effect sizes of other variables can be *overestimated* if the covariate is not reliably measured because the full effect of the covariate(s) would not be removed. This is a significant concern if the goal of research is to accurately model the "real" relationships evident in the population. Although most authors assume that reliability estimates (Cronbach alphas) of .7 to .8 are acceptable (e.g., Nunnally, 1978), and Osborne, Christensen, and Gunter (2001) report that the average alpha reported in top educational psychology journals is .83, measurement of this quality still contains

enough measurement error to make correction worthwhile, as illustrated below.

Correction for low reliability is simple, and widely disseminated in most texts on regression, but rarely seen in the literature. We argue that authors should correct for low reliability to obtain a more accurate picture of the "true" relationship in the population and, in the case of multiple regression or partial correlation, to avoid over-estimating the effect of another variable.

Since "the presence of measurement errors in behavioral research is the rule rather than the exception" and "reliabilities of many measures used in the behavioral sciences are, at best, moderate" (Pedhazur, 1997, p. 172), it is important that researchers be aware of accepted methods of dealing with this issue. For simple regression,

$$r_{12}^{\bullet} = \frac{r_{12}}{\sqrt{r_{11} r_{12}}}$$

provides an estimate of the "true" relationship between the IV and DV in the population. In this equation, r_{12} is the observed correlation, and r_{11} and r_{22} are the reliability estimates of the variables.

Even in cases in which reliability is .80, correction for attenuation substantially changes the effect size. When reliability drops to .70 or below, this correction yields a substantially different picture of the "true" nature of the relationship and potentially avoids a Type II error. With each independent variable added to the regression equation, the effects of less-than-perfect reliability on the strength of the relationship become more complex, and the results of the analysis become more questionable.

Homoscedasticity Assumption

Homoscedasticity means that the variance of errors is the same across all levels of the IV. When the variance of errors differs at different values of the IV, heteroscedasticity is indicated. According to Berry and Feldman (1985), slight heteroscedasticity has little effect on significance tests; however, when heteroscedasticity is marked, it can lead to serious distortion of findings and seriously weaken the analysis, thus increasing the possibility of a Type I error.

This assumption can be checked by visual examination of a plot of the standardized residuals (the errors) by the regression standardized predicted value. Most modern statistical packages include this as an option.

Ideally, residuals are randomly scattered around 0 (the horizontal line), providing a relatively even distribution. Heteroscedasticity is indicated when the residuals are not evenly scattered around the line. There are many forms heteroscedasticity can take, such as a bow-tie or fan shape. When the plot of residuals appears to deviate substantially from normal, more formal tests for heteroscedasticity should be performed. Possible tests for this are the Goldfeld-Quandt test when the error term either decreases or increases consistently as the value of the DV increases as shown in the fan-shaped plot, or the Glejser tests for heteroscedasticity when the error term has small variances at central observations and larger variance at the extremes of the observations as in the bow tie-shaped plot (Berry & Feldman, 1985). In cases where skew is present in the IVs, transformation of variables can reduce the heteroscedasticity.

Conclusion

The goal of this Digest was to raise awareness of the importance of checking assumptions in simple and multiple regression. We focused on four assumptions that were not highly robust to violations, or easily dealt with through design of the study, that researchers could easily check and deal with, and that, in our opinion, appear to carry substantial benefits.

We believe that checking these assumptions carries significant benefits for the researcher. Making sure an analysis meets the associated assumptions helps avoid Type I and II errors. Attending to issues such as attenuation due to low reliability, curvilinearity, and non-normality often boosts effect sizes, usually a desirable outcome.

Finally, there are many nonparametric statistical techniques available to researchers when the assumptions of a parametric statistical technique are not met. Although these are often somewhat lower in power than parametric techniques, they provide valuable alternatives, and researchers should be familiar with them.

References

Berry, W. D., & Feldman, S. (1985). *Multiple Regression in Practice.* Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07-050. Newbury Park, CA: Sage Publications, Inc.

Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Goldfeld, S.M. and Quandt, R.E. (1976). *Studies in Nonlinear Estimation.* Cambridge, MA: Ballinger Publishing Company.

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw Hill.

Osborne, J. W., Christensen, W. R., & Gunter, J. (April 2001). Educational psychology from a statistician's perspective: A review of the power and goodness of educational psychology research. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Pedhazur, E. J. (1997). *Multiple Regression in Behavioral Research* (3rd ed.). Orlando, FL: Harcourt Brace.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon.

BEST COPY AVAILABLE

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract ED99CO0032. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the U.S. Department of Education. **Permission is granted to copy and distribute this ERIC/AE Digest.**

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
X	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").